



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 10, 2024 – 08:00 am GMT

PDB ID : 4US5  
Title : Crystal Structure of apo-MsnO8  
Authors : Maier, S.; Pflueger, T.; Loesgen, S.; Asmus, K.; Broetz, E.; Paululat, T.;  
Zeeck, A.; Andrade, S.; Bechthold, A.  
Deposited on : 2014-07-03  
Resolution : 1.80 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.003 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.39

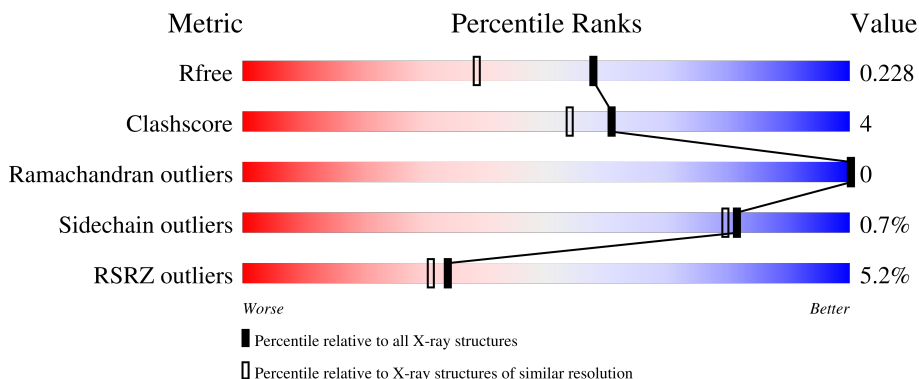
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

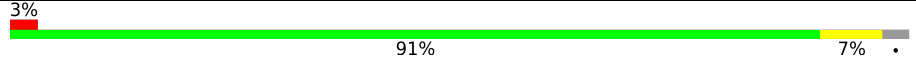
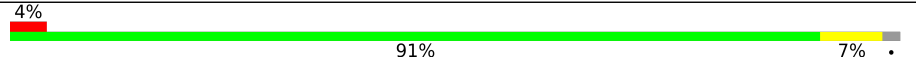
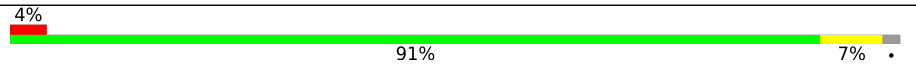
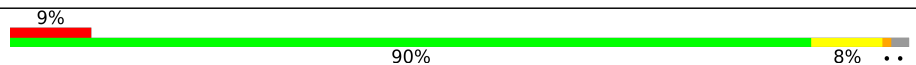
The reported resolution of this entry is 1.80 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	164625	7108 (1.80-1.80)
Clashscore	180529	8162 (1.80-1.80)
Ramachandran outliers	177936	8077 (1.80-1.80)
Sidechain outliers	177891	8076 (1.80-1.80)
RSRZ outliers	164620	7108 (1.80-1.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	338	 3% 91% 7% .
1	B	338	 4% 91% 7% .
1	C	338	 4% 91% 7% .
2	D	338	 9% 90% 8% ..

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 11466 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called LUCIFERASE-LIKE MONOOXYGENASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	N	O	S				Se
1	A	329	2589	1641	466	473	2	7	0	3	0
1	B	331	2612	1655	475	473	2	7	0	3	0
1	C	332	2637	1670	479	479	2	7	0	5	0

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	HIS	-	expression tag	UNP W8QCH3
A	-2	HIS	-	expression tag	UNP W8QCH3
A	-1	HIS	-	expression tag	UNP W8QCH3
A	0	HIS	-	expression tag	UNP W8QCH3
A	1	VAL	-	expression tag	UNP W8QCH3
B	-3	HIS	-	expression tag	UNP W8QCH3
B	-2	HIS	-	expression tag	UNP W8QCH3
B	-1	HIS	-	expression tag	UNP W8QCH3
B	0	HIS	-	expression tag	UNP W8QCH3
B	1	VAL	-	expression tag	UNP W8QCH3
C	-3	HIS	-	expression tag	UNP W8QCH3
C	-2	HIS	-	expression tag	UNP W8QCH3
C	-1	HIS	-	expression tag	UNP W8QCH3
C	0	HIS	-	expression tag	UNP W8QCH3
C	1	VAL	-	expression tag	UNP W8QCH3

- Molecule 2 is a protein called LUCIFERASE-LIKE MONOOXYGENASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	N	O	S				Se
2	D	331	2602	1649	471	473	2	7	0	2	0

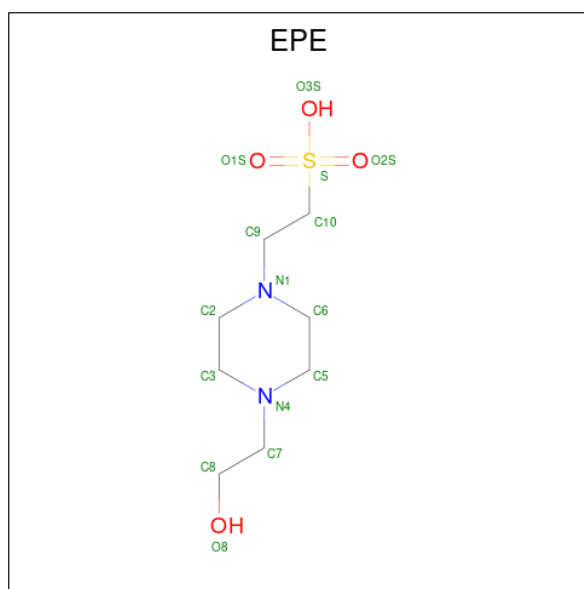
There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	-3	HIS	-	expression tag	UNP W8QCH3
D	-2	HIS	-	expression tag	UNP W8QCH3
D	-1	HIS	-	expression tag	UNP W8QCH3
D	0	HIS	-	expression tag	UNP W8QCH3
D	1	VAL	-	expression tag	UNP W8QCH3
D	175	ASP	GLU	conflict	UNP W8QCH3
D	193	ASP	GLN	conflict	UNP W8QCH3

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	6	Total Na 6 6	0	0
3	B	7	Total Na 7 7	0	0
3	C	7	Total Na 7 7	0	0
3	D	6	Total Na 6 6	0	0

- Molecule 4 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: C<sub>8</sub>H<sub>18</sub>N<sub>2</sub>O<sub>4</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N O S 15 8 2 4 1	0	0

*Continued on next page...*

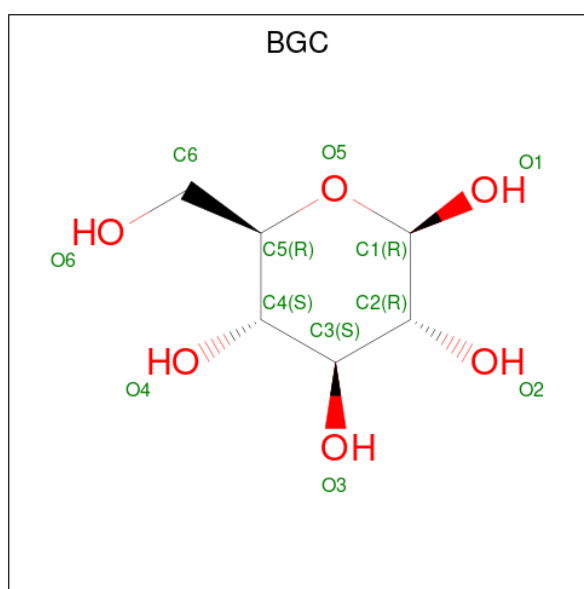
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	B	1	Total	C	N	O	S	0	0
			15	8	2	4	1		
4	C	1	Total	C	N	O	S	0	0
			15	8	2	4	1		
4	D	1	Total	C	N	O	S	0	0
			15	8	2	4	1		

- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

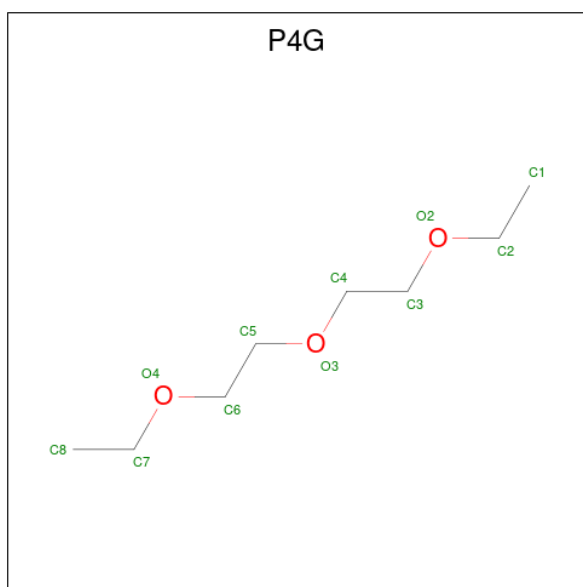
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	C	1	Total	Cl	0	0
			1	1		

- Molecule 6 is beta-D-glucopyranose (three-letter code: BGC) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	C	1	Total	C	O	0	0
			12	6	6		

- Molecule 7 is 1-ETHOXY-2-(2-ETHOXYETHOXY)ETHANE (three-letter code: P4G) (formula: C<sub>8</sub>H<sub>18</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	C	1	Total C O 11 8 3	0	0
7	C	1	Total C O 11 8 3	0	0
7	D	1	Total C O 11 8 3	0	0

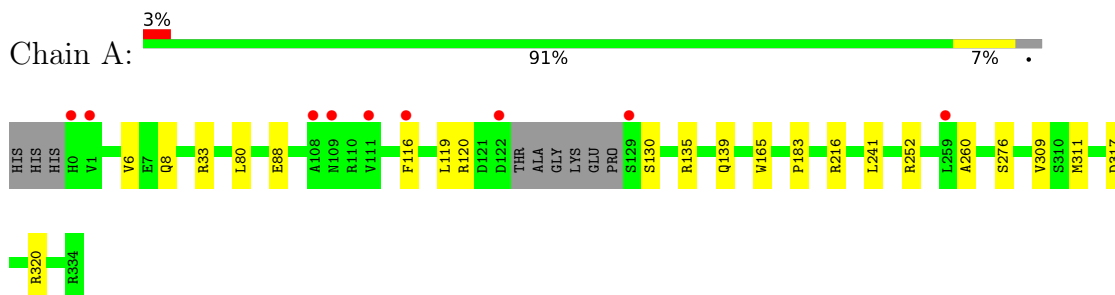
- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	277	Total O 277 277	0	0
8	B	267	Total O 267 267	0	0
8	C	205	Total O 205 205	0	0
8	D	145	Total O 145 145	0	0

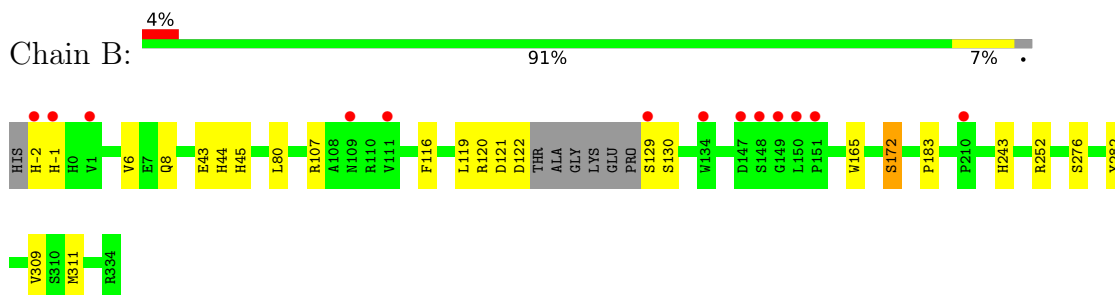
### 3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

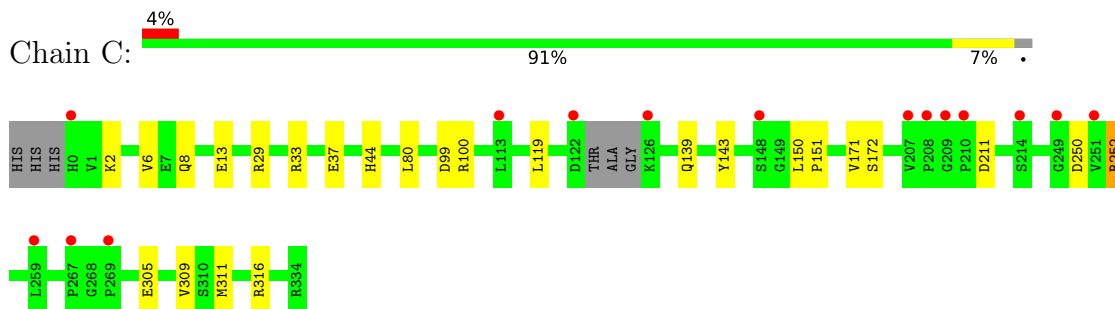
- Molecule 1: LUCIFERASE-LIKE MONOOXYGENASE



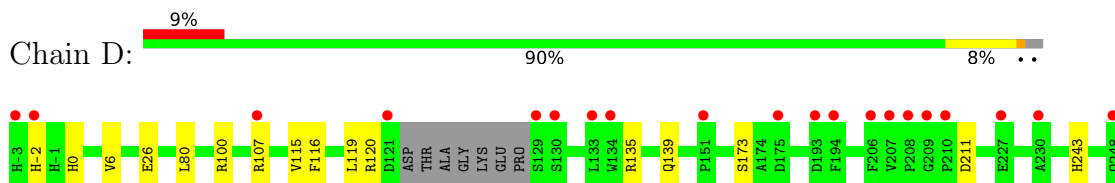
- Molecule 1: LUCIFERASE-LIKE MONOOXYGENASE

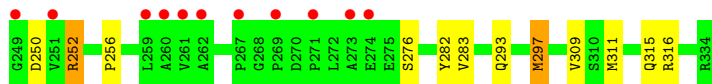


- Molecule 1: LUCIFERASE-LIKE MONOOXYGENASE



- Molecule 2: LUCIFERASE-LIKE MONOOXYGENASE







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	85.00Å 102.90Å 160.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.05 – 1.80 49.05 – 1.80	Depositor EDS
% Data completeness (in resolution range)	100.0 (49.05-1.80) 100.0 (49.05-1.80)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.50 (at 1.79Å)	Xtrriage
Refinement program	REFMAC 5.7.0032	Depositor
R, $R_{free}$	0.188 , 0.216 0.201 , 0.228	Depositor DCC
$R_{free}$ test set	6584 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	26.8	Xtrriage
Anisotropy	0.043	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 26.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	11466	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.34% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: BGC, NA, P4G, EPE, CL

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.68	0/2650	0.68	2/3600 (0.1%)
1	B	0.66	0/2675	0.66	1/3633 (0.0%)
1	C	0.63	0/2699	0.67	2/3664 (0.1%)
2	D	0.58	0/2663	0.67	3/3619 (0.1%)
All	All	0.64	0/10687	0.67	8/14516 (0.1%)

There are no bond length outliers.

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	252	ARG	NE-CZ-NH1	5.80	123.20	120.30
1	C	252	ARG	NE-CZ-NH2	-5.71	117.45	120.30
2	D	252	ARG	NE-CZ-NH2	-5.70	117.45	120.30
2	D	297	MSE	CG-SE-CE	5.57	111.15	98.90
2	D	0	HIS	CB-CA-C	-5.51	99.38	110.40
1	A	252	ARG	NE-CZ-NH2	-5.34	117.63	120.30
1	C	252	ARG	NE-CZ-NH1	5.14	122.87	120.30
1	B	252	ARG	NE-CZ-NH1	5.10	122.85	120.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2589	0	2577	16	0
1	B	2612	0	2600	20	0
1	C	2637	0	2631	27	0
2	D	2602	0	2579	28	0
3	A	6	0	0	0	0
3	B	7	0	0	0	0
3	C	7	0	0	0	0
3	D	6	0	0	0	0
4	A	15	0	18	1	0
4	B	15	0	18	0	0
4	C	15	0	18	1	0
4	D	15	0	18	0	0
5	C	1	0	0	0	0
6	C	12	0	12	0	0
7	C	22	0	36	6	0
7	D	11	0	18	2	0
8	A	277	0	0	2	0
8	B	267	0	0	8	0
8	C	205	0	0	7	0
8	D	145	0	0	4	0
All	All	11466	0	10525	86	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 4.

All (86) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:293:GLN:HG3	2:D:297:MSE:CE	1.55	1.36
2:D:293:GLN:HG3	2:D:297:MSE:HE2	1.19	1.14
1:A:320:ARG:HG2	2:D:-2:HIS:HD2	1.15	1.07
1:C:37:GLU:OE2	8:C:2035:HOH:O	1.80	0.99
2:D:293:GLN:CG	2:D:297:MSE:HE2	1.92	0.98
1:A:320:ARG:HG2	2:D:-2:HIS:CD2	2.02	0.94
2:D:293:GLN:CG	2:D:297:MSE:CE	2.47	0.92
2:D:293:GLN:HG3	2:D:297:MSE:HE3	1.51	0.90
1:C:311[A]:MSE:SE	8:C:2005:HOH:O	2.45	0.83
1:A:311[B]:MSE:SE	8:A:2006:HOH:O	2.46	0.83
1:A:317:ASP:OD1	2:D:-2:HIS:ND1	2.20	0.75
1:A:216:ARG:NH1	4:A:1341:EPE:O2S	2.20	0.73
1:C:171:VAL:H	7:C:1345:P4G:H12	1.54	0.72
1:B:172:SER:OG	8:B:2146:HOH:O	2.09	0.70

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:80:LEU:HD13	2:D:119:LEU:HD13	1.73	0.70
1:B:311[B]:MSE:SE	8:B:2017:HOH:O	2.61	0.67
1:A:88:GLU:OE2	1:B:45:HIS:HD2	1.78	0.67
1:B:43:GLU:OE2	1:B:45:HIS:HE1	1.78	0.66
2:D:311[B]:MSE:SE	8:D:2010:HOH:O	2.63	0.66
2:D:250:ASP:OD1	2:D:252:ARG:HD2	1.95	0.66
1:C:250:ASP:OD1	1:C:252:ARG:HD2	1.96	0.65
1:C:150:LEU:HG	1:C:151:PRO:HD2	1.81	0.61
2:D:283:VAL:CG2	2:D:297:MSE:HE1	2.31	0.61
2:D:283:VAL:HG21	2:D:297:MSE:HE1	1.83	0.60
1:C:250:ASP:OD1	1:C:252:ARG:CD	2.50	0.60
8:B:2258:HOH:O	1:C:33:ARG:NH2	2.35	0.59
1:C:139[A]:GLN:HG3	1:C:143:TYR:CE2	2.38	0.59
1:C:37:GLU:HG2	8:C:2035:HOH:O	2.02	0.59
2:D:250:ASP:OD1	2:D:252:ARG:CD	2.50	0.58
1:C:8:GLN:HE21	1:C:311[A]:MSE:SE	2.38	0.57
2:D:115:VAL:HG12	2:D:119:LEU:HD12	1.86	0.57
2:D:107:ARG:NH1	2:D:173:SER:HB3	2.20	0.56
1:C:33:ARG:HD3	8:C:2024:HOH:O	2.04	0.56
1:B:107:ARG:NH2	1:B:130:SER:HA	2.19	0.56
1:A:8:GLN:HE21	1:A:311[B]:MSE:SE	2.40	0.55
1:A:320:ARG:CG	2:D:-2:HIS:HD2	2.03	0.55
1:C:99:ASP:OD2	4:C:1343:EPE:H51	2.07	0.55
1:B:44:HIS:HE1	8:B:2056:HOH:O	1.91	0.54
1:B:44:HIS:HD2	8:B:2061:HOH:O	1.92	0.52
1:B:8:GLN:HE21	1:B:311[B]:MSE:SE	2.42	0.52
1:C:13:GLU:O	8:C:2009:HOH:O	2.19	0.52
1:B:-2:HIS:CA	8:B:2004:HOH:O	2.58	0.51
1:B:43:GLU:OE2	1:B:45:HIS:CE1	2.61	0.51
1:C:172:SER:HB3	7:C:1345:P4G:H31	1.93	0.51
1:C:171:VAL:H	7:C:1345:P4G:C1	2.22	0.50
2:D:26:GLU:OE1	8:D:2021:HOH:O	2.19	0.49
1:A:116:PHE:CE1	1:A:120:ARG:HG3	2.48	0.49
1:B:116:PHE:CE1	1:B:120:ARG:HG3	2.48	0.48
2:D:316:ARG:NH1	8:D:2136:HOH:O	2.46	0.48
1:B:-2:HIS:HA	8:B:2004:HOH:O	2.14	0.48
1:C:8:GLN:NE2	1:C:311[A]:MSE:SE	2.97	0.48
1:B:243:HIS:CE1	1:B:282:TYR:OH	2.67	0.48
2:D:135:ARG:NH1	2:D:139[A]:GLN:OE1	2.46	0.48
1:C:6:VAL:HB	1:C:309:VAL:HG22	1.97	0.47
2:D:6:VAL:HB	2:D:309:VAL:HG22	1.96	0.47

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:44:HIS:HE1	8:C:2039:HOH:O	1.98	0.47
1:A:8:GLN:NE2	1:A:311[B]:MSE:SE	2.98	0.47
1:C:100:ARG:HH11	7:C:1346:P4G:H62	1.80	0.46
1:C:172:SER:OG	7:C:1345:P4G:C4	2.63	0.46
1:A:135:ARG:NH1	1:A:139:GLN:OE1	2.47	0.46
2:D:100:ARG:HD3	7:D:1342:P4G:H22	1.98	0.46
1:B:8:GLN:NE2	1:B:311[B]:MSE:SE	2.99	0.46
1:B:121:ASP:O	1:B:122:ASP:HB2	2.16	0.46
1:B:243:HIS:HE1	1:B:282:TYR:OH	1.99	0.45
2:D:293:GLN:CD	2:D:297:MSE:HE2	2.36	0.45
1:B:6:VAL:HB	1:B:309:VAL:HG22	1.98	0.45
1:C:37:GLU:CG	8:C:2035:HOH:O	2.65	0.44
1:A:6:VAL:HB	1:A:309:VAL:HG22	2.00	0.44
1:A:165:TRP:CD2	1:A:183:PRO:HB2	2.53	0.43
1:B:80:LEU:HD13	1:B:119:LEU:HD13	2.00	0.43
1:A:80:LEU:HD13	1:A:119:LEU:HD13	2.01	0.43
1:A:33:ARG:NH2	8:A:2047:HOH:O	2.46	0.43
2:D:243:HIS:CE1	2:D:282:TYR:OH	2.72	0.43
2:D:116:PHE:CZ	2:D:120:ARG:HD2	2.54	0.42
1:B:-2:HIS:CD2	1:B:-1:HIS:H	2.37	0.42
1:C:29[A]:ARG:O	1:C:33:ARG:HG3	2.19	0.42
1:C:172:SER:OG	7:C:1345:P4G:H42	2.19	0.42
2:D:243:HIS:HE1	2:D:282:TYR:OH	2.02	0.42
1:C:2:LYS:HB2	1:C:305:GLU:HG2	2.01	0.42
1:B:165:TRP:CD2	1:B:183:PRO:HB2	2.55	0.41
8:B:2006:HOH:O	1:C:316:ARG:HD2	2.20	0.41
2:D:315:GLN:NE2	8:D:2018:HOH:O	2.35	0.41
1:C:80:LEU:HD13	1:C:119:LEU:HD13	2.03	0.41
1:C:211:ASP:OD1	1:C:211:ASP:N	2.54	0.40
1:A:241:LEU:HD22	1:A:260:ALA:HA	2.04	0.40
2:D:100:ARG:HH11	7:D:1342:P4G:C1	2.33	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	328/338 (97%)	320 (98%)	8 (2%)	0	100	100
1	B	330/338 (98%)	324 (98%)	6 (2%)	0	100	100
1	C	333/338 (98%)	325 (98%)	8 (2%)	0	100	100
2	D	329/338 (97%)	323 (98%)	6 (2%)	0	100	100
All	All	1320/1352 (98%)	1292 (98%)	28 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	272/270 (101%)	270 (99%)	2 (1%)	81	79
1	B	274/270 (102%)	271 (99%)	3 (1%)	70	65
1	C	277/270 (103%)	277 (100%)	0	100	100
2	D	273/270 (101%)	270 (99%)	3 (1%)	70	65
All	All	1096/1080 (102%)	1088 (99%)	8 (1%)	81	79

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	130	SER
1	A	276	SER
1	B	129	SER
1	B	172	SER
1	B	276	SER
2	D	211	ASP
2	D	256	PRO
2	D	276	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such

sidechains are listed below:

Mol	Chain	Res	Type
1	B	-2	HIS
1	B	0	HIS
1	B	44	HIS
1	B	45	HIS
1	B	243	HIS
2	D	243	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

### 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

### 5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

Of 35 ligands modelled in this entry, 27 are monoatomic - leaving 8 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# $ Z  > 2$	Counts	RMSZ	# $ Z  > 2$
7	P4G	C	1346	-	10,10,10	0.68	0	9,9,9	0.79	0
7	P4G	C	1345	-	10,10,10	0.62	0	9,9,9	0.70	0
4	EPE	B	1342	-	15,15,15	2.08	2 (13%)	18,20,20	2.63	8 (44%)
4	EPE	C	1343	-	15,15,15	1.65	2 (13%)	18,20,20	2.66	7 (38%)
7	P4G	D	1342	-	10,10,10	0.75	0	9,9,9	0.80	0
6	BGC	C	1344	-	12,12,12	0.72	0	17,17,17	0.96	1 (5%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	EPE	A	1341	-	15,15,15	1.96	2 (13%)	18,20,20	2.67	8 (44%)
4	EPE	D	1341	-	15,15,15	1.69	2 (13%)	18,20,20	2.58	7 (38%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	P4G	C	1346	-	-	4/8/8/8	-
7	P4G	C	1345	-	-	3/8/8/8	-
4	EPE	B	1342	-	-	3/9/19/19	0/1/1/1
4	EPE	C	1343	-	-	5/9/19/19	0/1/1/1
7	P4G	D	1342	-	-	7/8/8/8	-
6	BGC	C	1344	-	-	2/2/22/22	0/1/1/1
4	EPE	A	1341	-	-	2/9/19/19	0/1/1/1
4	EPE	D	1341	-	-	3/9/19/19	0/1/1/1

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1341	EPE	C10-S	-6.84	1.67	1.77
4	B	1342	EPE	C10-S	-6.78	1.67	1.77
4	D	1341	EPE	C10-S	-5.37	1.69	1.77
4	C	1343	EPE	C10-S	-5.31	1.70	1.77
4	B	1342	EPE	O2S-S	4.00	1.56	1.45
4	D	1341	EPE	O2S-S	3.21	1.54	1.45
4	C	1343	EPE	O1S-S	2.86	1.53	1.45
4	A	1341	EPE	O3S-S	2.11	1.55	1.47

All (31) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	1343	EPE	C6-N1-C2	6.14	122.66	108.83
4	B	1342	EPE	C5-N4-C3	5.63	121.51	108.83
4	B	1342	EPE	C6-N1-C2	5.35	120.86	108.83
4	A	1341	EPE	C6-N1-C2	5.26	120.67	108.83
4	D	1341	EPE	C5-N4-C3	5.21	120.56	108.83
4	A	1341	EPE	C5-N4-C3	5.15	120.43	108.83
4	D	1341	EPE	C6-N1-C2	5.06	120.22	108.83

*Continued on next page...*



Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	1343	EPE	C5-N4-C3	4.98	120.03	108.83
4	A	1341	EPE	C9-N1-C6	4.71	123.28	111.23
4	B	1342	EPE	C9-N1-C6	4.49	122.72	111.23
4	A	1341	EPE	C7-N4-C3	4.26	122.12	111.23
4	D	1341	EPE	C7-N4-C3	4.26	122.12	111.23
4	D	1341	EPE	C9-N1-C6	4.24	122.08	111.23
4	C	1343	EPE	C7-N4-C3	3.80	120.95	111.23
4	B	1342	EPE	C7-N4-C3	3.73	120.77	111.23
4	C	1343	EPE	O3S-S-C10	3.43	111.32	105.77
4	C	1343	EPE	C9-N1-C2	3.24	119.52	111.23
4	C	1343	EPE	C7-N4-C5	3.03	118.99	111.23
4	A	1341	EPE	O1S-S-C10	-2.72	103.63	106.92
4	A	1341	EPE	O3S-S-C10	2.66	110.07	105.77
4	C	1343	EPE	C9-N1-C6	2.58	117.83	111.23
4	D	1341	EPE	O2S-S-C10	2.54	109.98	106.92
4	B	1342	EPE	C7-N4-C5	2.54	117.72	111.23
4	D	1341	EPE	C9-N1-C2	2.52	117.67	111.23
4	A	1341	EPE	C7-N4-C5	2.39	117.35	111.23
4	B	1342	EPE	O2S-S-C10	2.39	109.79	106.92
4	D	1341	EPE	C7-N4-C5	2.36	117.28	111.23
6	C	1344	BGC	O2-C2-C3	-2.26	105.12	110.35
4	A	1341	EPE	O2S-S-C10	2.05	109.39	106.92
4	B	1342	EPE	C2-C3-N4	-2.02	106.50	110.64
4	B	1342	EPE	C9-N1-C2	2.02	116.39	111.23

There are no chirality outliers.

All (29) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1341	EPE	C10-C9-N1-C6
4	A	1341	EPE	C8-C7-N4-C3
4	B	1342	EPE	C10-C9-N1-C6
4	B	1342	EPE	C8-C7-N4-C3
4	C	1343	EPE	C8-C7-N4-C3
4	C	1343	EPE	C9-C10-S-O2S
4	C	1343	EPE	C9-C10-S-O3S
4	D	1341	EPE	C10-C9-N1-C6
4	D	1341	EPE	C8-C7-N4-C3
7	C	1346	P4G	C5-C6-O4-C7
6	C	1344	BGC	C4-C5-C6-O6
7	D	1342	P4G	O3-C5-C6-O4
6	C	1344	BGC	O5-C5-C6-O6

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
7	C	1346	P4G	O2-C3-C4-O3
7	C	1345	P4G	O2-C3-C4-O3
4	C	1343	EPE	C10-C9-N1-C2
4	D	1341	EPE	C10-C9-N1-C2
7	D	1342	P4G	C8-C7-O4-C6
7	D	1342	P4G	C4-C3-O2-C2
7	C	1345	P4G	O3-C5-C6-O4
7	C	1345	P4G	C5-C6-O4-C7
4	C	1343	EPE	C9-C10-S-O1S
7	D	1342	P4G	C3-C4-O3-C5
7	C	1346	P4G	C6-C5-O3-C4
4	B	1342	EPE	N4-C7-C8-O8
7	D	1342	P4G	C1-C2-O2-C3
7	D	1342	P4G	C5-C6-O4-C7
7	D	1342	P4G	O2-C3-C4-O3
7	C	1346	P4G	O3-C5-C6-O4

There are no ring outliers.

5 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	C	1346	P4G	1	0
7	C	1345	P4G	5	0
4	C	1343	EPE	1	0
7	D	1342	P4G	2	0
4	A	1341	EPE	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	323/338 (95%)	-0.13	9 (2%) 55 53	14, 25, 46, 85	2 (0%)
1	B	325/338 (96%)	-0.21	13 (4%) 43 40	12, 22, 49, 81	2 (0%)
1	C	326/338 (96%)	0.12	15 (4%) 38 35	11, 29, 63, 94	4 (1%)
2	D	325/338 (96%)	0.69	31 (9%) 15 13	20, 37, 67, 94	1 (0%)
All	All	1299/1352 (96%)	0.12	68 (5%) 34 31	11, 28, 60, 94	9 (0%)

All (68) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	-2	HIS	5.4
1	B	149	GLY	4.1
2	D	-3	HIS	4.1
1	A	111	VAL	4.1
2	D	261	VAL	3.6
2	D	273	ALA	3.6
1	A	129	SER	3.5
2	D	129	SER	3.4
1	C	207	VAL	3.4
1	A	109	ASN	3.3
2	D	134	TRP	3.3
1	B	129	SER	3.2
2	D	210	PRO	3.2
2	D	260	ALA	3.2
2	D	269	PRO	3.2
1	C	126	LYS	3.0
1	C	208	PRO	3.0
1	C	209	GLY	3.0
2	D	207	VAL	2.9
2	D	209	GLY	2.9
2	D	249	GLY	2.8

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	C	210	PRO	2.8
1	B	147	ASP	2.7
2	D	251	VAL	2.7
1	C	214	SER	2.7
1	B	-2	HIS	2.7
1	B	109	ASN	2.7
2	D	208	PRO	2.7
1	A	116	PHE	2.7
1	B	210	PRO	2.7
1	C	251	VAL	2.6
2	D	259	LEU	2.6
1	A	108	ALA	2.6
2	D	227	GLU	2.6
1	B	-1	HIS	2.5
1	A	0	HIS	2.5
1	A	122	ASP	2.5
2	D	248	GLN	2.5
1	B	151	PRO	2.5
1	C	259	LEU	2.4
2	D	121	ASP	2.4
2	D	107	ARG	2.4
1	C	0	HIS	2.4
1	B	150	LEU	2.4
2	D	267	PRO	2.4
1	B	111	VAL	2.4
1	C	249	GLY	2.3
1	A	1	VAL	2.3
2	D	193	ASP	2.3
1	C	148	SER	2.3
2	D	133	LEU	2.3
2	D	206	PHE	2.2
1	B	148	SER	2.2
1	C	113	LEU	2.2
1	B	134	TRP	2.2
2	D	274	GLU	2.1
1	B	1	VAL	2.1
2	D	194	PHE	2.1
1	C	269	PRO	2.1
1	C	122	ASP	2.1
2	D	230	ALA	2.1
2	D	130	SER	2.1
2	D	151	PRO	2.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	D	271	PRO	2.0
2	D	175	ASP	2.0
1	C	267	PRO	2.0
2	D	262	ALA	2.0
1	A	259	LEU	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
7	P4G	D	1342	11/11	0.54	0.28	48,55,65,66	0
7	P4G	C	1346	11/11	0.56	0.26	40,52,69,70	0
7	P4G	C	1345	11/11	0.81	0.19	25,48,62,64	0
6	BGC	C	1344	12/12	0.83	0.12	39,46,51,58	0
4	EPE	C	1343	15/15	0.86	0.16	34,56,69,73	0
4	EPE	D	1341	15/15	0.86	0.16	39,51,63,70	0
4	EPE	A	1341	15/15	0.86	0.16	34,50,66,70	0
3	NA	D	1336	1/1	0.89	0.13	40,40,40,40	0
4	EPE	B	1342	15/15	0.89	0.15	30,45,60,60	0
3	NA	C	1340	1/1	0.93	0.18	45,45,45,45	0
3	NA	D	1340	1/1	0.93	0.12	41,41,41,41	0
3	NA	D	1335	1/1	0.94	0.12	39,39,39,39	0
3	NA	B	1341	1/1	0.95	0.10	31,31,31,31	0
3	NA	C	1336	1/1	0.95	0.12	36,36,36,36	0
3	NA	D	1337	1/1	0.95	0.16	44,44,44,44	0
3	NA	B	1337	1/1	0.95	0.08	35,35,35,35	0
3	NA	C	1341	1/1	0.96	0.08	34,34,34,34	0
3	NA	A	1339	1/1	0.96	0.13	39,39,39,39	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	NA	D	1339	1/1	0.96	0.10	36,36,36,36	0
3	NA	C	1335	1/1	0.97	0.09	31,31,31,31	0
3	NA	D	1338	1/1	0.97	0.15	36,36,36,36	0
3	NA	B	1340	1/1	0.97	0.12	34,34,34,34	0
3	NA	A	1337	1/1	0.97	0.06	33,33,33,33	0
3	NA	A	1336	1/1	0.98	0.10	31,31,31,31	0
3	NA	B	1339	1/1	0.98	0.07	28,28,28,28	0
3	NA	C	1337	1/1	0.98	0.15	29,29,29,29	0
3	NA	C	1338	1/1	0.98	0.10	30,30,30,30	0
3	NA	B	1336	1/1	0.99	0.07	24,24,24,24	0
3	NA	A	1338	1/1	0.99	0.05	26,26,26,26	0
3	NA	B	1338	1/1	0.99	0.05	27,27,27,27	0
5	CL	C	1342	1/1	0.99	0.03	31,31,31,31	0
3	NA	A	1335	1/1	0.99	0.05	28,28,28,28	0
3	NA	C	1339	1/1	0.99	0.13	30,30,30,30	0
3	NA	A	1340	1/1	0.99	0.06	32,32,32,32	0
3	NA	B	1335	1/1	0.99	0.07	26,26,26,26	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.